"VIDEO ENCODING AND DECODING METHODS AND CORRESPONDING ENCODING AND DECODING DEVICES"

FIELD OF THE INVENTION

The present invention generally relates to the field of video compression and, for instance, to the video coding standards of the MPEG family (MPEG-1, MPEG-2, MPEG-4) and to the video recommendations of the ITU-H.26X family (H.261, H.263 and extensions, H.264). More specifically, this invention relates to an encoding method applied to an input video sequence corresponding to successive scenes subdivided into successive video object planes (VOPs) and generating, for coding all the video objects of said scenes, a coded bitstream the content of which is described in terms of separate channels and constituted of encoded video data in which each data item is described by means of a bitstream syntax allowing to recognize and decode all the elements of said content, said syntax comprising an additional syntactic information provided for describing independently the type of temporal prediction of the various channels, said predictions being chosen within a list comprising the following situations :

- the temporal prediction is formed by directly applying the motion field sent by the encoder on one or more reference pictures ;
- the temporal prediction is a copy of a reference image ;
- the temporal prediction is formed by the temporal interpolation of the motion field ;
- the temporal prediction is formed by the temporal interpolation of the current motion field and further refined by the motion field sent by the encoder .

The invention also relates to a corresponding encoding device, to a transmittable video signal consisting of a coded bitstream generated by such an encoding device, and to a method and a device for decoding a video signal consisting of such a coded bitstream.

BACKGROUND OF THE INVENTION

In the first video coding standards and recommendations (up to MPEG-4 and H.264), the video was assumed to be rectangular and to be described in terms of a luminance channel and two chrominance channels. With MPEG-4, an additional channel carrying shape information has been introduced. Two modes are available to

compress those channels : the INTRA mode, according to which each channel is encoded by exploiting the spatial redundancy of the pixels in a given channel of a single image, and the INTER mode, that exploits the temporal redundancy between separate images. The INTER mode relies on a motion-compensation technique, which describes an image from one or several image(s) previously decoded by encoding the motion of pixels from one image to the other. Usually, the image to be encoded is partitioned into independent blocks or macroblocks, each of them being assigned a motion vector. A prediction of the image is then constructed by displacing pixel blocks from the reference image(s) according to the set of motion vectors (luminance and chrominance channels share the same motion description). Finally, the difference (called the residual signal) between the image to be encoded and its motion-compensated prediction is encoded in the INTER mode to further refine the decoded image. However, the fact that all pixel channels are described by the same motion information is a limitation damaging the compression efficiency of the video coding system.

SUMMARY OF THE INVENTION

It is therefore the object of the invention to propose a video encoding method in which said drawback is avoided by adapting the way the temporal prediction is formed.

To this end, the invention relates to a method such as defined in the introductory part of the description and which is moreover characterized in that said additional syntactic information is a syntactic element placed in said generated coded bitstream and its meaning is specific for each present channel, said element being placed at the slice level or at the macroblock level according to the proposed embodiment.

The invention also relates to a corresponding encoding device, to a transmittable video signal consisting of a coded bitstream generated by such an encoding device, and to a method and a device for decoding a video signal consisting of such a coded bitstream.

DETAILED DESCRIPTION OF THE INVENTION

According to the invention, it is proposed to introduce in the encoding syntax used by the video standards and recommendations an additional information consisting of a new syntactic element supporting their lack of flexibility and opening new possibilities to encode more efficiently and independently the temporal prediction of

various channels. This additional syntactic element, called for example "channel temporal prediction", takes the following symbolic values :

Motion_compensation

Temporal_copy

Temporal_interpolation

Motion_compensated_temporal_interpolation,

and the meaning of these values is :

a) motion_compensation : the temporal prediction is formed by directly applying the motion field sent by the encoder on one or more reference pictures (this default mode is implicitly the INTER coding mode of most of the current coding systems) ;

b) temporal_copy : the temporal prediction is a copy of a reference image ;

c) temporal_interpolation : the temporal prediction is formed by the temporal interpolation of the motion fields ;

d) motion_compensated_temporal_interpolation : the temporal prediction is formed by the temporal interpolation of the current motion field and further refined by the motion field sent by the encoder.

The words "temporal interpolation" must be understood in a broad sense, i.e. as meaning any operation of the type defined by an expression such as Vnew = a.V1 + b.V2 + K, where V1 and V2 designate previously decoded motion fields, a and b designate coefficients respectively assigned to said motion fields, K designates an offset and Vnew is the new motion field thus obtained. It can therefore be seen that, in fact, the particular case of the temporal copy is included in the more general case of the temporal interpolation, for b = 0 and K = 0 (or a = 0 and K = 0).

According to the invention, the additional syntactic element thus proposed has to be placed at the following levels in the coded bitstream that has to be stored (or to be transmitted to the decoding side):

1) either at the slice level ;

2) or at the macroblock level ;

this additional syntactic element being in each case either specific for each present channel or, possibly, shared by all the channels.

This invention may be used in some identified situations where the way of constructing the temporal prediction can be switched on a slice or macroblock basis, and also on a channel basis. A first example may be for instance a sequence with a shape channel : it is possible that the shape information does not change much, whereas the

luminance and chrominance channels carry varying information (it is for instance the case with a video depicting a rotating planet : the shape is always a disc, but the texture of it depends on the planet rotation). In this situation, the shape channel can be recovered by temporal copy, and the luminance and chrominance channels by motion compensated temporal interpolation. A second example may be the case of a change at the macroblock level. In a video sequence showing a seascape with the sky in the upper part of the picture, unlike the sea, the sky remains the same from one image to the other. Its macroblocks can therefore be encoded by temporal copy, whereas the macroblocks of the sea have to be encoded by temporal interpolation.